DOCUMENT RESUME

ED 441 296 EC 307 811

AUTHOR Stinson, Michael; Eisenberg, Sandy; Horn, Christy; Larson, Judy; Levitt, Harry; Stuckless, Ross

TITLE Real-Time Speech-to-Text Services. [A Report of the] National Task Force on Quality of Services in the Postsecondary Education of Deaf and Hard of Hearing Students.

INSTITUTION National Technical Inst. for the Deaf, Rochester, NY.

SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC.; Association on Higher Education and Disability, Columbus, OH.

PUB DATE 1999-00-00

NOTE 27p.; In collaboration with the Postsecondary Education Programs Network (PEPNet) and co-sponsored by the Conference of Educational Administrators of Schools and Programs for the Deaf (CEASD).

AVAILABLE FROM Northeast Technical Assistance Center, Rochester Institute of Technology, National Technical Institute for the Deaf, 52 Lomb Memorial Dr., Rochester, NY 14623-5604. Tel: 716-475-6433 (Voice/TTY); Fax: 716-475-7660.

PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS *Assistive Devices (for Disabled); Classroom Communication; *Communication Aids (for Disabled); *Computer Assisted Instruction; *Computer Mediated Communication; *Deafness; *Hearing Impairments; Higher Education

IDENTIFIERS *Transcription

ABSTRACT

This report describes and discusses several applications of new computer-based technologies which enable postsecondary students with deafness or hearing impairments to read the text of the language being spoken by the instructor and fellow students virtually in real time. Two current speech-to-text options are described: (1) steno-based systems in which a stenographer uses a 24-key machine to encode spoken English phonetically into a computer where it is converted into English text and displayed on a computer screen; and (2) computer-assisted notetaking in which a typist with special training uses a standard keyboard to input words into a laptop/PC as they are being spoken. Both types of systems provide a real-time text output that students can read on a computer or television screen in order to follow what is occurring in class. For each of these systems, the report addresses how the system works, major considerations that need to be addressed with respect to its implementation as a support service in the classroom, who is qualified to provide the service, training needs, and how the system's effectiveness can be evaluated. An appendix contains a form for evaluating student needs for real-time stenographic services. (Contains 37 references.) (CR)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 441 296

National Task Force on Quality of Services
in the Postsecondary Education of
Deaf and Hard of Hearing Students

Report on
REAL-TIME SPEECH-TO-TEXT SERVICES

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

EC 307811

BEST COPY AVAILABLE

National Task Force on Quality of Services in the
Postsecondary Education of Deaf and Hard of Hearing Students

REAL-TIME SPEECH-TO-TEXT SERVICES

AUTHORS:

Michael Stinson
National Technical Institute for the Deaf

Sandy Eisenberg
California State University, Northridge

Christy Horn
University of Nebraska

Judy Larson
St. Louis Community College

Harry Levitt
City University of New York

Ross Stuckless
National Technical Institute for the Deaf

EDITOR AND TASK FORCE CHAIR:
Ross Stuckless

NETAC COPY EDITOR:
Kathleen Smith
Rochester Institute of Technology
National Technical Institute for the Deaf
Rochester, New York

1999

Suggested citation:
Stinson, M., Eisenberg, S., Horn, C., Larson, J., Levitt, H., Stuckless, R.
Real-Time Speech-to-Text Services: A report of the National Task Force on Quality of Services in the Postsecondary Education of Deaf and Hard of Hearing Students.
Rochester, N.Y.: Northeast Technical Assistance Center, Rochester Institute of Technology.

Distributed by the Northeast Technical Assistance Center (NETAC) in collaboration with the Postsecondary Education Programs Network (PEPNet) and co-sponsored by the Office of Special Education and Rehabilitative Services (OSERS) of the U.S. Department of Education (84.078A); the Association on Higher Education and Disability (AHEAD); and the Conference of Educational Administrators of Schools and Programs for the Deaf (CEASD).

Rochester Institute of Technology

National Technical Institute for the Deaf
Northeast Technical Assistance Center
52 Lomb Memorial Drive
Rochester, NY 14623-5604
716-475-6433 (V/TTY)
716-475-7660 (Fax)

BEST COPY AVAILABLE

Editor's note

This is one in a series of reports intended to assist postsecondary institutions in developing and maintaining special services of quality as needed by their deaf and hard of hearing students. Each report has been prepared with postsecondary administrators, faculty, and staff uppermost in mind, and particularly those most likely to have a role in providing services to these students. It is anticipated that these reports will be useful also to deaf and hard of hearing students in gaining more information about services for which they may be eligible.

A challenge in authoring and editing each of these reports is to avoid giving the impression that all the information they contain pertains equally to all deaf and hard of hearing students at the postsecondary level. Of course this is not so. These students are individuals first, and their needs and wishes for special services and other accommodations will vary, as will characteristics of the particular colleges and universities they as individuals choose to attend.

Also, it is a challenge to write about needs and services for both deaf and hard of hearing students together. While they do share a hearing loss, the magnitude of their hearing loss ranges collectively from mild to profound. But while the special needs of deaf students may be more apparent than those of hard of hearing students, the special needs of hard of hearing students are no less real.

Twelve reports are scheduled for distribution in 1997-99, each with a different focus and each authored by a working committee of experts on a particular subject. All are members of a **National Task Force on Quality of Services in the Postsecondary Education of Deaf and Hard of Hearing Students**. This task force was formed in 1994 and numbers 100 members associated with 32 two and four-year colleges in 28 states and provinces in the United States and Canada.

Readers are free to cite information and views from each of the reports and to duplicate and share copies. In return, they are asked to cite the names of its authors and make bibliographic reference to the report.

– Ross Stuckless

REAL-TIME SPEECH-TO-TEXT SERVICES

Michael Stinson, Sandy Eisenberg, Christy Horn, Judy Larson, Harry Levitt, and Ross Stuckless[1]

INTRODUCTION

Real-time speech-to-text has been defined as the accurate transcription of words that make up spoken language into text momentarily after their utterance (Stuckless, 1994).

This report will describe and discuss several applications of new computer-based technologies, which enable deaf and hard of hearing students to read the text of the language being spoken by the instructor and fellow students, virtually in real time. In its various technological forms, real-time speech-to-text is a growing classroom option for these students.

This report is intended to complement several other such reports in this series which focus on notetaking (Hastings, Brecklein, Cermack, Reynolds, Rosen, & Wilson, 1997)[2], assistive listening devices (Warick, Clark, Dancer, & Sinclair, 1997), and interpreting (Sanderson, Siple, & Lyons, 1999). It is notable that the Department of Justice has interpreted the Americans with Disabilities Act (P.L. 101-336) to include computer-aided transcription services under "appropriate auxiliary aids and services" (28CFR, §36.303).

It should be emphasized at the outset that the real-time speech-to-text services described and discussed in this report are intended to complement, not replace, the options that are already available.

Development of real-time speech-to-text systems

Over the past 20 years, several developments have made it possible to use real-time speech-to-text transcription services as we know them today. These began with the development of smaller, more powerful computer systems, including their capability of converting stenotypic phonetic abbreviations electronically into understandable words. These parallel developments led to the earliest applications of steno-based systems both to the classroom and to real-time captioning in 1982.

In the later 1980s, laptop computers became widely available. This enhanced portability led to the use of computers for notetaking in which the notetaker used a standard keyboard in the regular classroom. It was at this time that stenotype machines were also linked to laptop computers, enhancing their portability. In the late 1980s, abbreviation software became available for regular keyboards (Stinson & Stuckless, 1998).

Currently, both steno-based and standard keyboard approaches are being used with deaf and hard of hearing students in many mainstream secondary and postsecondary settings. Although the full extent of their usage nationwide remains to be documented, over the past 10 years there clearly has been an increased demand for speech-to-print transcription services in the classroom (Cuddihy, Fisher, Gordon, & Shumaker, 1994; Haydu & Patterson, 1990; James & Hammersley, 1993; McKee, Stinson, Everhart, & Henderson, 1995; Messerly & Youdelman, 1994; Moore, Bolesky, & Bervinchak, 1994; Smith & Rittenhouse, 1990; Stinson, Stuckless, Henderson, & Miller, 1988; Virvan, 1991).

Two current speech-to-text options

Currently, two major options are available for providing real-time speech-to-text services to deaf and hard of hearing students. The first and second parts of this report will discuss these two options in order. But first, several general comments about the two systems should be made.

Steno-based systems. For these systems, a trained stenographer uses a 24-key machine to encode

---

[1] In the order listed above, the authors are associated with National Technical Institute for the Deaf (Rochester, New York), California State University, Northridge (Northridge, California), University of Nebraska (Lincoln, Nebraska), St Louis Community College (St. Louis, Missouri), City University of New York (New York, NY), and National Technical Institute for the Deaf.

[2] The report on notetaking made reference also to computer-assisted notetaking, C-Print ™, and real-time captioning, each of which is an application of real-time speech-to-text. In the present report, frequent reference is made to the generation of notes as a secondary application of real-time speech-to-text.

spoken English phonetically into a computer where it is converted into English text and displayed on a computer screen or television monitor in real time. Generally, the text is produced verbatim. When used in schools, this system is often called CART (computer-aided real-time transcription), an apt acronym in view of the fact that stenotypists often transport their equipment from one classroom to another on wheels.

Computer-assisted notetaking systems. For these systems, a typist with special training uses a standard keyboard to input words into a laptop/PC as they are being spoken. Sometimes these take the form of summary notes, sometimes almost as verbatim text. These systems are often abbreviated as CAN (computer-assisted notetaking).

Both types of systems provide a real-time text output that students can read on a computer or television screen in order to follow what is occurring in class. In addition, the text file can be examined by students, tutors, and instructors after class either on the screen or as hard copy.

These technologies offer receptive communication to deaf and hard of hearing students. However, they provide limited options for expressive communication on the part of these students, and service providers need to keep this in mind.

We will begin by providing some basic "nuts and bolts" information that service providers need in order to implement a steno-based or computer-assisted notetaking (CAN) system. For each of these systems, we address four major questions:

(1) How do these systems work?
(2) What major considerations need to be addressed with respect to their implementation as a support service in the classroom?
(3) Who is qualified to provide the service, and what is his/her training?
(4) How can the system's effectiveness be evaluated, and what has been learned from evaluations to date?

In considering these systems, we will discuss aspects of particular speech-to-text systems with which we have had personal experience. Our focus on particular systems or associated college programs is not intended as an endorsement over other systems or college programs.

The third part of this report pertains to the use of speech-to-print services relative to other forms of support service, and the fourth part to the development of new speech-to-text systems, focusing on the status and potential of automatic speech recognition (ASR).

Steno-Based Systems

Steno-based systems began to be used in classrooms in 1982, with mainstreamed deaf and hard of hearing students at Rochester Institute of Technology (Stuckless, 1983). Today, steno-based systems rank as an effective support service for large numbers of deaf and hard of hearing students in mainstream college environments throughout the country. This growth is due to a number of factors, including refinements in the necessary software; faster, more reliable, and more portable computers; the increasing availability of stenographic reporters (and in many cases the lowering cost of their services); and most important, generally favorable classroom evaluations (Stinson, Stuckless, Henderson, & Miller, 1988).

How steno-based systems work

The person who provides this service in the educational setting may be called a stenotypist, stenographer, or stenographic or educational reporter. His/her equipment typically includes a laptop with several cables and special software, a stenographic machine that has been designed to interface with the laptop and its software, and a display of some kind for presenting the student with the text.

The stenotypist can display the text in real time in several ways, using a TV or computer monitor (including the screen of a second connected laptop), or projecting the text onto a screen by using an LCD or overhead projector. Unlike conventional captioning, which superimposes a line or two of text over a picture, real-time steno-generated text can fill a full screen. Depending on the need, the text output of a steno-based system can be displayed in the classroom itself and/or elsewhere via electronic connections.

Typically the stenotypist is present in the classroom with the deaf or hard of hearing student. However, depending on his/her level of skill and familiarity with the topic under discussion, it is also possible to

use a phone link to transmit speech to a stenotypist in a distant location, returning the text to the student via a second telephone line or using a computer modem. Cellular phones have also been used successfully for this purpose where fixed telephone lines were not available (Kanevsky, Nahamoo, Walls, & Levitt, 1992).

Equipment. The equipment consists of three basic components: a computer-compatible stenographic machine, an IBM-compatible laptop, and the software needed to convert the stenographic input of speech and display it as text.

Stenographic machine. The stenographic machine, similar to that used by "computer-connected" court reporters, permits the stenotypist to "write" (key in) verbatim dialogue at speeds of 200 wpm or greater.[3] These speeds are possible in large part because he/she can "chord" keys, depressing several keys simultaneously instead of sequentially as in conventional typing.

Laptop computer. A Pentium 166 MHz or faster laptop, with at least 32 MB of memory and an active-matrix screen is recommended. Two serial ports are preferred, but a PCMCIA slot is acceptable.

Software. The translation software is the heart of the system. Several companies produce the software, and each stenotypist has his/her favorite. Among the most popular are RapidText (Irvine, CA) and Cheetah Systems (Fremont, CA). Essentially, the software consists of four parts, often incorporated into a single software product.

(1) large built-in dictionary (50,000 words or larger), with provisions for the stenotypist to make additions as new words arise in class,
(2) program which selects words from the dictionary based on a specific logic and set of rules, and
(3) word processing program that arranges these words in a particular format and performs other editing tasks.

The following chart shows examples of steno code and their corresponding English words.

Steno code (input)	English text (output)
WREUG	writing
O	on
-T	the
PH-PB/APS	machine's
KAOE/PWORD	keyboard

(4) encoding software to format and display the text in tandem with any of several peripheral devices, e.g., TV or computer monitor, laptop screen, projected image, or printer/paper copy.

Need for technical support. It should be emphasized that a steno-based system is a technologically sophisticated service. Software needs to be installed correctly, and hardware needs to be set up properly. Students depend on the system, and if it breaks down it will need to be repaired promptly, so technical support should be available and close at hand.

Applications with deaf and hard of hearing students

Steno-based systems provide a two-fold service that includes real-time speech-to-text transcription for deaf and hard of hearing students to read almost instantly in the classroom, and a written record of the class that they can use later for review. We will discuss these two applications in turn.

Real-time classroom implementation. Steno-based systems can be used to cover a variety of campus events, sometimes as "real-time captioning" where the text appears under the video image of a speaker. However, their primary application with deaf and hard of hearing students is in the classroom. Steno-based systems as used in the regular classroom provide a means for the deaf or hard of hearing student to replace listening with reading what the teacher and fellow students are discussing, in near real time.

As indicated earlier, the stenotypist sits near the front of the classroom, sometimes to the side where he/she is in visual range of the teacher, students, the chalkboard, and other visual media that might be in use. Incidentally, the stenotypist's equipment is silent and requires little space.

So long as the text is legible to the deaf or hard of hearing student, it can be displayed in a number of ways. If the service is being provided for a single student, a second laptop can be used as a screen. However, if a number of deaf and/or hard of hearing students are using the service, a large TV or projection screen is in order.

[3] Parenthetically, the average speaking rate of college teachers as they lecture is around 150 words per minute, with a standard deviation across the faculty of about 30 wpm (Stuckless, 1994).

From a classroom perspective, the presence of a steno-based system or a computer-assisted notetaking system in the class is similar in some respects to having an interpreter there. More attention will be given to similarities and differences later in this report.

Hard copy text. Transcripts of lectures can be used as complete classroom notes, preserving the entire lecture and all students' comments for subsequent review by deaf and hard of hearing students taking the course. Typically, these transcripts are shared with these students and with the instructor. Some instructors welcome the transcripts as a way of tightening their lectures and reviewing their students' questions and comments.

If the instructor chooses, he/she should be at liberty to share them with hearing members of the class also.[4] The transcripts can be of value also in tutoring deaf and hard of hearing students, enabling tutors to organize tutoring sessions in close accord with course content. Also, interpreters sometimes use them to improve their signing of course-specific words and expressions.

Once the stenotypist has completed the real-time transcription of a class for the deaf or hard of hearing student(s) enrolled in the course, he/she will edit the text. Depending on the particular class, a 50-minute class is likely to generate 25 to 30 pages of text.

If the stenotypist has a high accuracy rate in a given class, e.g., 98-99%, he/she may be able to correct errors and make the text more readable in one-half hour or less. Obviously more errors (causes of which are discussed later under Accuracy) will require more editing time.

Many students who use the text for review purposes prefer receiving an ASKII disk (edited or unedited) so they can organize their own format and decide for themselves what they want to retain or discard.

Accuracy

The most important task for the stenotypist working in the classroom is to maintain high accuracy in the production of text from speech. When the accuracy drops below 95%, i.e., more than one word error in 25, intelligibility of the text drops off rapidly.[5]

The following excerpt from a lecture[6] illustrates some of the types of errors that can appear with steno-based systems. The upper line indicates what the teacher said, and the lower line indicates a transcribed text version.

(Speech) Interestingly enough one of the most popular courses
(Text) INTERESTINGLY ENOUGH ONE OF THE MOST POPULAR COURSES

on this campus is a course on *death* and dying. Since so many *of* us are
ON THIS CAMPUS IS A COURSE ON *DEARTH* AND DYING. SINCE SO MANY ___ US ARE

trying to avoid that I have some *ambivalent* feelings about the
TRYING TO AVOID THAT I HAVE SOME *ALL BEVELLENTD* FEELINGS ABOUT THE

popularity of that course. I do *know* that it's a very popular
POPULAR ARE THE OF THAT COURSE. I DO *NO* THAT IT'S A VERY POPULAR

course and at the same time I *know* that it's a subject that most of us
COURSE AND AT THE SAME TIME I *NO* THAT IT'S A SUBJECT THAT MOST OF US

want to avoid.
WANT TO AVOID.

Types of errors. Based on the number of departures in the text from what was spoken, there are six word errors in the above 65-word spoken excerpt, yielding 90% accuracy. We can see the four most common types of word errors illustrated in the text above:

mistranslate	–	*death/DEARTH, popularity/ POPULAR ARE THE*
omission	–	*of/ _*
untranslate	–	*ambivalent/ALL BEVELLENTD*
homonym	–	*know/NO (2)*

[4] It is common for stenographic reporters in private practice to add a surcharge for distribution of extra copies of the text. In the educational environment, this should be discouraged.
[5] This pertains to all the real-time speech-to-text systems discussed in this report.
[6] This particular lecture was given in February 1982 at NTID/ Rochester Institute of Technology, as part of the first course in which a steno-based system was ever used. Today, we look for better than 95% accuracy.

Sources of errors. There are at least three general sources of errors:

(a) *Stenotypist errors.* The computer is unforgiving of input errors on the part of the stenotypist. Once made, they cannot be corrected "online".

(b) *Vocabulary limitations.* Each stenotypist is expected to add and maintain his/her own special course-related dictionary of words beyond the large dictionary that comes with the software. The goal here is nearly perfect pre-edited text, so ongoing dictionary-building time (like editing time) should be built into the service.

The textbooks used in class should be made available to the stenotypist for the purpose of dictionary building. Instructors are also encouraged to share specialized vocabulary likely to be used in class with the stenotypist so he/she can enter this vocabulary into his/her dictionary prior to the class meeting. Over time, the accuracy of the stenotypist's work will improve as he/she builds a specialized dictionary and his/her stenotyping errors diminish.

(c) *Teacher/classroom/course content factors.* Some teachers and hearing classmates of the deaf and hard of hearing students articulate more clearly and/or speak more slowly and deliberately than others. Also, some are more grammatically "correct" in their speech than others.

Adverse classroom factors include "noisy" classroom conditions, e.g., several people speaking simultaneously. The stenotypist cannot be expected to produce meaningful, accurate text under these conditions.

By their very nature some areas of study lend themselves better to the use of steno-based systems than others. For example, courses demanding considerable physical activity and foreign language courses may be poor prospects for the use of steno-based systems.

The Stenotypist

Some stenotypists provide their services on an hourly basis, and some by the academic term. Still others are employed as members of the college's professional staff. Mostly this depends on the number and year-to-year continuity of deaf and hard of hearing students likely to be requesting the service.

A college with just one student requesting the service is unlikely to hire a stenotypist on a long-term basis when there is no assurance that the student will complete his/her program of studies in the same institution. At the other end, a college that has an ongoing need to provide steno-based services for numerous deaf and hard of hearing students each year is likely to prefer hiring stenotypists as regular staff members.

Training. The starting point for becoming a stenotypist is training in a stenographic or court-reporting school, of which there are more than 400 throughout the country. Many stenotypists and most active court reporters are affiliated with the National Court Reporters Association (NCRA). Both court reporting and stenotyping in the college setting require high-speed, accurate stenographic translation of the spoken word, often involving multiple speakers. Most court reporters, however, *ipso facto* are not adept at providing real-time transcription in the classroom. They have the luxury of being able to edit their material before producing a readable transcript.

In contrast, stenotypists in the classroom situation must produce near-perfect accuracy without the benefit of prior editing. This calls for special skills that overlap with those of real-time TV captionists and which come with training (if available) and experience. When feasible, it is useful for the beginning stenotypist to have a semester of practice time, and time to build his/her special dictionary, before taking on full responsibility for supporting students in the classroom. Another opportunity for practice is to produce transcripts from videotapes for captioning purposes.

Certification. The National Court Reporters Association offers certification at several levels. Some stenotypists argue that NCRA certification has little relevance to working as a stenotypist in the classroom, but certification undeniably provides added assurance of both speed and accuracy.[5]

Recruitment. Sometimes the most direct and efficient way to recruit stenotypists, at least for short term, temporary support, is through local stenographic agencies. Insist on real-time experience

[5] The Center on Deafness at California State University Northridge periodically offers workshops for stenotypists interested in working with deaf and hard of hearing students attending college.

and require that they provide their own hardware and software (including their own dictionary).

Local court reporting/stenographic schools may be able to provide leads from among their own graduating students and graduates. For long-term recruitment of stenotypists into college programs for deaf and hard of hearing students, an internship agreement with one of these schools can be an effective way of incorporating newly graduated real-time stenotypists into the college's support services for deaf and hard of hearing students.

Pay levels. Compensation standards for stenotypists working with deaf and hard of hearing students at the college level vary considerably, based on training and experience. Colleges with little or no prior experience using real-time stenotypists in the classroom may wish to check with other colleges that have, before varying much either way from the following ranges.

For "educational realtime reporters" with full-time (two semester, 40 hour week) college positions, the National Court Reporters Foundation (NCRF) of NCRA has suggested a salary range of $20,000 to $38,500 plus a full benefits package.[8] This range can be adjusted for use in colleges that use another calendar such as the quarter.

For those who are retained on an hourly fee basis, NCRF has suggested the following: $40-$75 per class hour (2-hour minimum), $15-$40 per hour for preparation time (30 minutes for each class hour), and $15-$40 per hour for production time (editing for distribution). However, fees of up to $150 per hour have been reported.

The importance of preparation and editing has already been discussed. Typically those who provide the service on an hourly fee basis furnish their own steno machines, laptops, and software.

Workloads. On-line classroom stenotyping requires sustained and undivided attention. And like teaching and interpreting, when done without periodic breaks it can be mentally and physically fatiguing. As a rule, for full-time staff, course coverage should not exceed 20-22 class hours per week. Back-to-back classes should be infrequent. Between-class time, e.g., three to four hours a day, can be used mainly for preparation and editing purposes. First-time coverage of new courses (and different instructors teaching the same courses) will require more preparation and editing time than those previously covered.

Evaluation of the service. Support service providers need some way to determine whether students using a steno-based system are being adequately served. Two aspects of evaluation are (a) quality of the real-time display and the hard-copy text, and (b) student/consumer feedback regarding his/her benefits from use of the system.

Quality of real-time display and edited text. Early and later on in the course, the stenotypist's college supervisor should appraise the quality of the display and the edited text for each course being covered by the stenotypist. The supervisor's principal interest here is that the real-time display be relatively free of errors (recognizing that the stenotypist is not the source of all errors), and that its format contribute to its readability. This can be determined by examining the ***unedited*** text, including word correctness/errors, punctuation, paragraphing, and indications of changes in speakers.

The *edited* text should be appraised relative to its intelligibility and ease of student use for review purposes.

Student/consumer feedback. Students using the steno-based service should be asked to make a formal evaluation midway through the course. Information may be collected on the student's perceptions regarding the skills and attitudes of the stenotypist. The Appendix shows a sample form used at California State University, Northridge to obtain student/consumer feedback.

In addition, each instructor who uses the steno system in his/her class should be given the opportunity to express his/her perception of the value of the service relative to its use by the deaf or hard of hearing student(s) in the class.

A study conducted with deaf and hard of hearing students at Rochester Institute of Technology taking courses in the College of Business and/or Liberal

---

[8] Information from *Realtime in the educational setting: Implementing new technology for access and ADA Compliance* (1994), National Court Reporters Foundation: Vienna, VA. Booklet available through NCRA Member Services and Information Center, 8224 Old Courthouse Rd., Vienna, VA 22182-3808.

Arts indicated that students responded favorably to the system, although there was variability in their responses. A majority of the students reported that they understood more from the steno-based text display than from interpreting (Stinson, Stuckless, Henderson, & Miller, 1988).

When supporting an individual student, a steno-based or other speech-to-print system obviously is more expensive when combined with other services such as interpreting, than when it is the only service provided, i.e., used "stand alone". It may be difficult to justify the provision of both the speech-to-text service and interpreting services for a single student in the class.

Nor do there appear to be consistent policies for dealing with such requests in colleges around the country when one student in the class requests speech-to-text, and another requests interpreting. In some circumstances, both services have been provided, whereas in others, students have been limited to only one of these services. Clear guidelines regarding when to provide one or both services remain to be developed.

A caveat on steno-based systems

In the hands of competent stenotypists, steno-based real-time speech-to-text offers a powerful support service to many deaf and hard of hearing students in college. Unfortunately, the relatively high costs of well-qualified stenotypists (not their equipment), together with their scarcity in most locations of the country, combine to make the service unavailable or underused in many colleges.

With this in mind, we proceed to examine some related alternatives.

Computer-Assisted Notetaking (CAN): Computer Systems with Standard Keyboards

When used with deaf and hard of hearing students, computer-assisted notetaking (CAN) systems, like steno-based systems, are used primarily in the classroom, in lieu of interpreters and notetakers. Like steno-based systems, CAN converts speech into text in real time for the deaf or hard of hearing student to read in the classroom. And like steno-based systems, CAN provides the student with an edited or unedited copy of the text for use as notes.

Unlike steno-based systems, CAN involves the use of a standard keyboard and a typist with special training, referred to in this report as a captionist but called a transcriber in some settings. There are a number of CAN systems, each of which varies in its details. In general, these systems all involve a (hearing) captionist sitting in the classroom and using a standard keyboard and a commercially available word processing program (such as WordPerfect) to transcribe information as it is being spoken in class.

The text is displayed in real time for deaf and hard of hearing students to read on a TV monitor or on a second laptop display (depending upon the number of deaf or hard of hearing students using that system in the particular class). At the end of class, the text is saved as a word processing file that can then be edited, printed, and distributed to these students as hard copy.

Keyboard input

Various CAN systems have "evolved" from the use of standard typing (character by character). The limitation of standard typing, even at high speed, is that it cannot keep up with the speed of speaking. Instructors' speaking rates typically run around 150 words per minute, and sometimes in bursts exceeding 200 words per minute.

Nevertheless, one basic approach is simply to substitute the handwriting of notes (at around 30 words per minute) with typing (at around 60 words per minute) – that is, the typist takes down in summary what the instructor says. With the advent of laptop "notebook" computers, this has become common among students who take notes for themselves, and increasingly among those who take notes for deaf and hard of hearing students (Hastings, Brecklein, Cermak, Reynolds, Rosen, & Wilson, 1997).

Various CAN systems employ different strategies to enable the captionist to increase his/her speed of input in order to capture more spoken content and detail. The goal is to come as close as possible to capturing all the relevant information being discussed in class, in a readable format. Two strategies are employed to enable transcribers to cover as much information as possible: (a) computerized abbreviation systems to reduce keystrokes, and (b) text-condensing strategies to enable the transcriber to type fewer words without losing spoken information.

Cost and personnel advantages over steno-based systems

CAN systems have several practical advantages over steno-based systems. CAN systems use portable, low-cost equipment. Also, the potential pool of typists/captionists is much larger than that of stenotypists and the costs of their services are usually lower than those of well-qualified stenotypists or interpreters. In general, the special training required for a well-qualified typist to become an acceptable CAN captionist can be a month or less, depending on the specific goals of the system (McKee et al., 1995).

Several CAN systems have been developed for or used in providing support services to deaf and hard of hearing students. The following table presents a summary of characteristics of eight computer-assisted systems for which published information is available.

Summary of characteristics of different computer-assisted notetaking systems

Characteristics of systems

System	*Uses abbreviation to increase speed*	*Location of text display*	*Attempts verbatim or real-time notes*	*Communication between student and transcriber*	*Required skills and/or training*
CAN-Cleveland (Messerley & Youdelman, 1994)	Not described	Connect with monitor	Generally, summary notes	One-way communication: transcriber to student	Minimal: must be able to summarize, type more than 60 wpm; good English use
Computer-Assisted Notetaking-NY (Kozma-Spytek & Balcke, 1995)	Not described	Connect with laptop	Summary notes	Two-way communication between transcriber and student	None described
CAN-Washington, D.C. (Virvan, 1991)	Abbreviations used as long as everyone understands. Does not use computerized abbreviation expansion program	Connect with monitor or laptop	Usually summary notes, but is capable of near verbatim transcription	One-way communication: transcriber to student	Overall little training, but required skills are ability to type over 60 wpm and summarize well, good English skills, hear well
C-Note (Cuddihy, Fisher, Gordon, & Shumaker, 1994)	Student & transcriber develop appropriate shorthand system	Connect with laptop	Varies from near-verbatim to summary notes	Two-way communication between transcriber and student	Not described
Project CONNECT (Knox Quinn & Anderson-Inman, 1996)	Not described	Connect with laptop	Summary notes and near-verbatim text	Two-way communication between transcriber and student	Not described
C-Print™ NTID System (McKee, Stinson, Everhart, & Henderson, 1995; Everhart, Stinson, McKee, Henderson, & Giles, 1996)	Emphasizes extensive use of phonetically-based abbreviation system to reduce key strokes	Connect with monitor or laptop	Near-verbatim text	Two-way communication between transcriber and student	Transcriber should be able to type 60 wpm. Formal course provided, with 62 page training manual and 50 training audiotapes
InstaCap (Hobeland, 1988; Warick, 1994)	Uses single key strokes to invoke full words for 20 abbreviations	Wireless connection with monitor	Varies from near-verbatim to summary	One-way communication: transcriber to student	Not fully described; transcriber's skills evaluated every 3-5 years
Notebook Computer Notetaking System (James & Hemmesley, 1993)	Not described	Connect with laptop	Generally summary notes	Two-way communication between transcriber and student	Not described

BEST COPY AVAILABLE

Hardware

The hardware used for CAN systems is simpler than that required for steno-based systems. However, when used in tandem with appropriate software, it can be sufficient to produce an effective text display.

Laptop computer. The basic piece of equipment is a laptop computer. Some systems use IBM-compatible computers (e.g., IBM's ThinkPad, NEC Versa 2000). Others report using Apple Macintosh PowerBooks (Messerley &Youdelman, 1994).

Display. The real-time text on the transcriber's laptop can be displayed for the deaf or hard of hearing student using (a) a second laptop computer, (b) a VGA-to-TV adapter that connects the laptop to a regular TV monitor, or (c) an LCD projection display.

Software

A CAN system requires word processing software and in most instances communication software. The more sophisticated systems also use abbreviation software.

Word processing software. Products such as WordPerfect 6 and Word 97 often have special built-in features that increase their effectiveness, such as WordPerfect's "Macro" and "QuickCorrect" features. These permit creating the abbreviations of a limited number of words and phrases for input into a computer.

Communication software. This software permits communication between two or more laptop computers by creating an asynchronous link. These systems include C-Note (Cuddihy et al., 1994) and Carbon Copy (McKee et al., 1995).

This software provides two ways of communicating between two computers: (a) a full-screen mode, where only one individual can enter a message at a time, and (b) a split-screen mode where both individuals may enter messages simultaneously. Most of these programs permit scrolling back to review previous material on the student's computer while new material is being entered on the captionist's computer. (Cost: $200).

Word abbreviation software. Several software packages have been developed specifically for extensive abbreviation of words and phrases being entered into the computer. At this time, the two systems most commonly used with CAN appear to be the following:

Productivity Plus	*Instant Text*
Productivity Software International, Inc.	Textware Solutions
1220 Broadway	83 Cambridge St.
New York, NY 10001	Burlington, MA 01803-4181

Using one of these systems, the computer automatically converts the abbreviations typed by the captionist into the full words that appear on the screen. This software serves to increase typing speed without increasing the necessary number of keystrokes, and permits the text to more closely approach the speed of the talker.

An example of the application of one of these abbreviation systems to a CAN service is a speech-to-text transcription system called C-Print™ which was developed at the National Technical Institute for the Deaf (McKee, Stinson, Giles, Colwell, Hager, Nelson-Nasca, & MacDonald, 1998).[9] C-Print™ uses an extensive word-abbreviation dictionary, along with specific text-condensing strategies.

A major difference between C-Print™ and other CAN systems is its commitment to coming as close as possible to providing a verbatim transcription, due largely to the extensive abbreviation system it employs. As the teacher (or class participant) talks, the captionist types a series of abbreviations. For each abbreviation, Productivity Plus searches the dictionary for its equivalent full word and displays it on the screen. Two examples of abbreviations and their expansions as used in C-Print™ appear below.

Abbreviations	*Full expansions*
t kfe drqr	the coffee drinker
slvg t pblm	solving the problem

The C-Print™ captionist is not required to memorize all the abbreviations in the C-Print™ system. Instead, she/he learns a set of phonetic rules developed specifically for C-Print™, which are then applied to any English word that has been added to its system's general dictionary. The general

[9] The C-Print™ project has been supported by grants 180J3011 and 180U6004 from the United States Department of Education, Office of Special Education.

dictionary developed by the C-Print™ staff currently contains approximately 10,000 words, including suffixes, which were selected from research on word frequencies in English. Specialized dictionaries can also be created that allow for the abbreviation of vocabulary, phrases, and acronyms unique to a course or subject area.

Text display

Format. The text display for a CAN system generally shows words appearing letter by letter, as opposed to a steno-based system that displays individual words or groups of words in a single burst. For the C-Print™ system, the student sometimes sees a split-second conversion from the abbreviation to the full word. Student feedback indicates this is not distracting.

The number of lines of text displayed in real time varies by the type of display and size of letters. A single-spaced laptop display may show 30 or more lines of text. A television monitor display with letters of a large font size, such as 30-point, may permit up to 15 lines, depending upon the particular system.

Content. For the C-Print™ system, the operator does not type every word, but does try to capture as much important information as possible. The text generated by some CAN systems (for both real-time display and hard copy) can be considerably more detailed than notes taken by trained notetakers, but is more condensed than the transcriptions provided by steno-based systems. Below is an unedited paragraph of text, with follow-up comments, produced in a history class by a C-Print™ captionist. Note the use of complete sentences.

> *Professor*: King has successfully gone into Birmingham after the failure in Albany, and has provoked a great deal of violence and has gotten a great deal of press coverage. It is severe violence. Although violence is seen on national television and Kennedy responds by not defending the existing legislation as Eisenhower did, this is a crucial shift, but by saying he will create legislation in support of the cause. That is the Civil Rights Bill of June, 1963. He is initiating his own legislation. It would strengthen desegregation in all places. In response to this is the march on Washington that takes place on Aug. 28, 1963. This is in support of Kennedy's bill.
>
> Bayard Rustin and A. Philip Randolph come back into the picture to organize the event. King gives his famous "I have a dream" speech. It is a great symbolic event. It shows a great deal of unity within the country behind doing something about civil rights.
>
> *Student*: Is that an all-Black march?
>
> *Professor*: No. It was by no means an all-Black march, it was greatly diverse. A. Philip Randolph gets his dream of the march, but it is not all Black. The movement is unified around one strategy — provoke violence, get it on television, and get government to do something.

At the end of class, the CAN text is saved as a word processing file that can then be corrected and distributed to students as hard copy text, on a floppy disk, or electronically. Electronic distribution requires that the captionist have access to a computer and can send the file electronically to the student. The student in turn can download and print the text at his/her convenience. Student feedback indicates that an effort should be made to distribute the text on the same day as the class or the following day.

Preparation for class

The captionist has a number of duties prior to actual in-class transcription. In preparation for each class, she/he needs to become familiar with new terms and concepts likely to be used in class. If working with a CAN system that uses extensive abbreviations, she/he may add abbreviations to the specialized dictionary so that words used frequently in a particular course (e.g., technical words, proper names, new terms) will appear when their corresponding abbreviations are typed.

Equipment must be set up prior to the class. This may mean connecting two laptops with each other. If a television monitor is to be used, it must be requisitioned and connected.

Prior to the first class, the captionist should discuss with the students for whom the speech-to-text will be used how the CAN system works, what they can expect from it, and their respective responsibilities. They may also need to discuss specific ways in which the captionist can be helpful during class. This may include matters such as repeating the students'

questions if they're not understood in class, or reading aloud the questions and other comments the student types on his/her laptop with the intent of sharing them with the class. The latter assumes that the particular student chooses not to voice for him/herself, and that the particular CAN system being used has this interactive feature.

If the class activity is a small group discussion, it is desirable for the real-time display to be a laptop monitor rather than a television monitor. It seems easier for deaf and hard of hearing students to shift between viewing a laptop display directly in front of them and observing the speaker(s) than to shift attention between a television monitor and the speaker(s).

Preparation and distribution of notes[10]

The hard copy notes are intended to be educational tools, not necessarily near-verbatim accounts of what happened in class. Therefore, information that is extraneous to the educational content can be omitted. Also, any confidential information about the students or others should be omitted. The captionist should be sensitive to the wishes of the instructor regarding other information to be omitted from the hard copy notes for a particular class.

Assignments should be accurately recorded. Beyond assignments, a good approach for captionists to use when deciding what information to include and what to omit is to provide notes that would help a student who was absent know what educational information was presented. This approach will help captionists decide what to include, and what changes to make, to render the class content both accurate and understandable.

Ergonomics and the scheduling of captionists

Transcribing for more than one hour without a break increases the risk of what has variously been called repetitive motion injuries and cumulative trauma disorder. Captionists in the college environment are likely to engage in intense typing of continuous lectures for up to one hour and will generally need an hour of "down" time before resuming typing. This time can often be devoted to preparing notes or preparing for the next class.

In an attempt to minimize ergonomic risk factors, it is recommended that:

(a) captionists continue to develop their skills with the abbreviations system to reduce keystrokes, and use other text condensing strategies

(b) where possible, captionists choose seating that reduces discrepancies in table, elbow, and keyboard height

(c) regular interviews with the captionist be conducted by her/his supervisor to monitor changes in comfort, fatigue, and effort

(d) where feasible, the college make the captionist's position part time.

Qualifications and training of a CAN captionist

Qualified captionists need first to be skilled typists (with typing speeds of 60 words per minute or better), need to have good verbal and auditory skills, and need to be familiar with the operation of laptop computers. It is helpful if the captionist has familiarity with the course material, although this often is impractical as a requisite.

A survey of existing pay scales suggests an hourly rate ranging from $10 - $15, inclusive of preparation time and time required for text editing and distribution as notes. One college surveyed indicated a pay scale comparable to that of interpreters.

With respect to training, the C-Print™ system at NTID appears to be the only college offering CAN training as a course (McKee, Stinson, Everhart, & Henderson, 1995). This one-month course is designed to teach the abbreviation rules that enable the C-Print™ captionist to save substantial numbers of keystrokes. The course also teaches strategies to condense information. Training includes practice transcribing real college lectures from audiotapes. Training materials consist mostly of a 62-page manual and 50 audiotapes.

---

[10] This topic and several others that follow draw extensively from McKee, B., Stinson, M., Giles, P., Colwell, J., Hager, A., Nelson-Nasca, M., & MacDonald, A. (1998). *C-Print™: A Computerized Speech-to-Print Transcription System: A Guide for Implementing* C-Print™. Rochester, NY: National Technical Institute for the Deaf.

Regardless of the CAN system that is used, a real issue is how soon captionists can become comfortable displaying what they are typing in real time in the classroom. Coming into the classroom and keying in rapidly spoken lecture material, which will be viewed by a student who is dependent upon it for learning, is a challenging and sometimes stressful task.

Captionists may be concerned about keeping up with a lecture pace, omitting important information, and making errors. Before they can become comfortable doing this, they may need in-class experience transcribing lectures where the text is not displayed in real time for the student.

Illustrative Policies and Procedures

As with steno-based systems, the cooperation of the captionist, the deaf and hard of hearing students, hearing classmates, and the instructor is necessary in order for the CAN service to work successfully in the classroom. The following policies and procedures are adapted from those developed for one college (NTID) in its use of a CAN system (Giles, 1996), and are organized around *General Information, Captionist's Responsibilities, and Student's Responsibilities.*

General Information

- CAN notes are intended to be used by supported student(s) registered in the course and should not be copied unless otherwise specified by the instructor.
- CAN notes are not a substitute for attending class.
- Because the notes need to be edited quickly and distributed as soon as possible CAN notes are not guaranteed to have 100% correct grammar or spelling.

Captionist's Responsibilities

The captionist will:

- provide an in-class text display for appropriate support service students. In addition, notes (generated from the text display) will be made available to supported students who attended class.
- make every effort to type spoken information word-for-word, and communicate the information in the manner in which it is intended. At times (during fast speech), the captionist will need to summarize information, but she/he will type as much of the important information as possible.
- assist by voicing comments or questions typed by students on the laptop provided (if it has the necessary communication software), or in another way mutually agreed upon.
- begin typing upon arrival of the students. Any announcements made by the instructor before the student(s) arrive will be typed. After 10 minutes, if none of the supported students are in attendance, the captionist will leave. However, if the student has notified the CAN office or the instructor at least 24 hours in advance, the captionist will take notes if approved by the instructor.
- indicate different speakers in the text by indicating "Professor", "Female Student", and "Male Student".
- be responsible for facilitating communication between the supported student(s) and others, i.e., the instructor and other students. This includes asking for clarification from the instructor or other students when necessary.
- be responsible for trying to resolve any problems stemming from student or instructor concerns about CAN.
- arrive at least 10 minutes before class to allow time for equipment set up.
- become familiar with the scheduled lecture by preparing for class through reviewing the textbook and related materials.
- find a replacement if she/he is sick. If a replacement cannot be found, the captionist will notify the appropriate Support Department that will notify the supported student(s).
- provide on-the-spot troubleshooting for equipment breakdown with minimum disruption to the class. If no solution is found, the captionist will make an effort to accommodate the supported student(s) to the best of his/her ability. Technical breakdowns are unforeseen and most often require diagnoses outside the classroom environment.
- when necessary, request an interpreter for special circumstances such as an oral presentation by the supported student(s).
- provide class handouts to authorized individuals, e.g., tutors.
- Summarize videotapes (captioned or uncaptioned).

Student's Responsibilities

The student will:

- introduce him/herself to the captionist so the captionist is familiar with each student.
- be responsible for taking notes and diagrams from the blackboard and overhead.
- be responsible for notifying the CAN Office if he/she will not be attending class or has withdrawn from the course. Three consecutive unexcused absences will result in the termination of CAN services.
- be responsible for double-checking spelling on any vocabulary.
- raise her/his hand when interested in communicating comments or questions through typing on the laptop (if so equipped).
- inform the captionist of any special needs for special circumstances, e.g., interpreter, at least two weeks in advance.

Evaluating CAN services

In evaluating the effectiveness of CAN services, college staff will want to consider (a) the quality of the real-time display in class, and (b) the quality of the hard-copy text or notes distributed to students after class (together with the timeliness of their distribution). Evaluation should be tied to the objectives of the system, i.e., summary notes vs. near-verbatim text.

If the intent is that the captionist record as much information as possible, there is a need for some kind of comparison between what the teacher and students actually said in class and what the captionist typed. For example, some preliminary data indicate that it is possible for a CAN system to capture 65 percent of the total ideas expressed in a lecture and 83 percent of the important ideas. These figures were obtained by using a standardized procedure for comparing recordings of teachers' lecture material with the corresponding text typed by the captionists.

It is also important to obtain deaf and hard of hearing student feedback regarding (a) the benefit of the real-time display, (b) the extent of their understanding of the classroom discourse, (c) their ability to participate in class, (d) the professionalism of the captionist and appropriateness of her/his behavior, and (e) helpfulness of the notes.

Feedback should be obtained also from the captionist and the instructor. The evaluation form for stenotypists as shown in the Appendix can be modified for use in connection with CAN systems.

Questions for the instructor can include whether the role of the captionist was adequately explained, whether the captionist performed her/his job with minimum disruption to the class, whether teaching methods were altered to accommodate the CAN system, and whether the instructor was able to express her/his concerns to the captionist.

To date, the systematic collection of feedback regarding CAN systems from students and faculty has been limited. One major theme that emerges from all the reports is that students perceived these various systems as beneficial, particularly in creating increasing understanding of classroom communication (Hobelaid, 1988; McGee et al., 1995; Everhart, Stinson, McKee, & Giles, 1996).

Data also have been collected in the process of evaluating the C-Print™ system at Rochester Institute of Technology. Questionnaire interview data from mainstreamed deaf and hard of hearing students indicated that they reported significantly greater understanding of information during a lecture with C-Print™ than with an interpreter. In addition, students stated a preference for the hard-copy detailed notes generated by the C-Print™ system over notes from a traditional notetaker (Everhart et al., 1996).

These findings are similar to those for steno-based systems, but should not be construed to suggest that such systems should replace these more traditional services. The important point is that these data do show that some students and some classes find the services beneficial.

Relative advantages of Steno-based and CAN systems

Steno-based systems. Steno-based systems have the following advantages:

- Steno-based systems capture virtually every word that is spoken. Thus, it is possible for the student to read the text of exactly what was said in real time.
- One stenotypist can cover a two-hour class, with a brief break.
- The stenotype machine is virtually silent.

BEST COPY AVAILABLE

CAN systems. CAN systems have the following advantages:

- CAN systems yield notes that are briefer and potentially easier to study than the verbatim transcripts yielded by steno-based systems.
- CAN captionists require relatively little special keyboard training beyond the ability to type 60 words per minute, increasing their availability.

Consideration of the relative advantages of the two systems indicates that it is not possible to make a general recommendation of one system over the other. A college may even wish to include both services in its repertoire of technologies.

The decision regarding which of the two services to provide will depend on a variety of issues, including availability of potential staff to provide support, costs, the type of class, and individual student needs.

New Technologies for Communication among Machines

A relatively recent application of technology, used most often with steno-based systems, is the provision of real-time transcription between two "remote" sites by telephone lines. The voice of a speaker is picked up by a microphone and transmitted to a stenotypist at a remote location via the first of two telephone lines. The stenotypist relays the real-time text via a second telephone line back to a television or computer display for the deaf or hard of hearing individual to read where he/she is located (Preminger & Levitt, 1997; Eisenberg & Rosen, 1996; Levitt, 1994; Stuckless, 1994). Although reports of this approach describe applications only with steno systems, it should apply also with CAN systems.

Infrared and *radio frequency-based networking devices* use a technology that increases the portability and ease of use of speech-to-text systems in the classroom. This technology eliminates the need for the cables that are commonly used to connect laptop computers with each other. One drawback of these cables is that the two laptop computers, i.e., the one being used by the captionist or stenotypist and the one being used by the student, need to be relatively close to each other. Also, cable connections require set-up time (often between classes) and are inconvenient when strung out in a classroom setting.

Infrared networking devices use a PCMCIA adapter (such as *Cooperative* that is produced by *Photonics*), or devices now being integrated into many laptop models, permitting wireless communication between computers. This means that the two (or more) computers do not need to be in close proximity to each other, and time does not need to be devoted to connecting the computers (Knox-Quinn & Anderson-Inman, 1996).

Software that permits two-way communication between the student and captionist or stenotypist already has been described. Network software (such as *Aspects* produced by *Group Logic*), provides for real-time collaborative interaction among up to 32 persons working in the same word-processing or graphics document. This network software permits the stenotypist or captionist to simultaneously communicate with more than one other computer, i.e., with numerous students in different locations of the classroom.

Using this software, it is also possible to create a split-screen display in which students may communicate with each other or add their own notes on half the screen, while observing the CAN or steno-generated text on the other half (Knox-Quinn & Anderson-Inman, 1996). One particular benefit of such an arrangement is that it may encourage note-taking on the part of the deaf or hard of hearing student, since she/he need not look at the keyboard. An added feature is that the program can correlate the student's own notes with the CAN or steno-generated text.

Use of Real-Time Speech-to-Text Relative to Other Classroom Support Services

Real-time speech-to-text is one of four direct classroom support services that are discussed in this series of reports, the others consisting of *assistive listening devices*, *interpreting*, and *notetaking*. Some of the factors we should consider in choosing one or more of these services with a given deaf or hard of hearing student taking a particular course follow. These factors are classified loosely under *Individual deaf or hard of hearing student*, *Course and/or instructor*, and *Other considerations*. For the purpose of this report, we will discuss these only in relation to real-time speech-to-text services.

Factors to be considered in selection of real-time speech-to-text and alternative support services in the classroom

Individual deaf or hard of hearing student. Student-specific factors include:

- **Preference of the student.**
  Major consideration should be given to providing this service when it is the student's preference over other services.
- **Prior experience and satisfaction with specific classroom support service.**
  Favorable prior experiences in using real-time speech-to-text in the classroom support the student's preference.
- **Ability to participate orally in question-asking and discussion.**
  Real-time speech-to-text services require that students either use their own voice if their speech is intelligible, or type and have the captionist read the display aloud to the class. For students with intelligible speech, it generally is easier for them to speak than to type.
- **Ability to make effective use of an assistive listening device in the classroom.**
  If the student is able to make effective use of an assistive listening device in the classroom, if the device is well maintained, and if both the instructor and fellow students cooperate in its use, the student may have little need for the real-time service. However she/he may continue to need its notetaking features.
- **Level of reading proficiency.**
  A requisite for functional use of real-time speech-to-text at the college level is the student's ability to read the text.
- **Level of signing proficiency.**
  A deaf student is likely to have proficiency in sign language, and this may be her/his first language. If so, the student may profit more from the use of an interpreter than from real-time speech-to-text. However, this will not obviate the probable need for a notetaking service of some kind.

Course and/or instructor. Course/instructor factors include:

- **Lecture vs. discussion-oriented course.**
  Some courses involve more active in-class student participation than others. Because of the interactive constraints on real-time speech-to-text systems, they are better adapted to courses that feature a lecture mode than to courses that are highly discussion-oriented. This reservation may not apply to students with intelligible speech skills.
- **Course content.**
  In general, speech-to-print services may work less effectively with certain courses, such as mathematics. However, experience in providing services indicates that the student's preferences and needs are critical in deciding which of his/her courses should use speech-to-text services. Where one student may not feel that a computer science class is appropriate for speech-to-text services, another student may.
- **Duration of class period.**
  Regardless of the type of service, a class extending beyond an hour without a break can be stressful for the service provider. Given a 10-minute break after the first hour, the stenotypist providing a steno-based service appears to be better able to continue through the second hour without relief than the captionist offering a CAN service or the interpreter providing the interpreting service.
- **Instructor's communication style.**
  The perfect instructor for real-time speech-to-text services (and for interpreting and conventional notetaking services as well) is one who speaks at or below normal speaking rates, i.e., 150 wpm, articulates clearly, and tends to use grammatically correct sentence structures. She/he is well organized by topic, and shares her/his lecture notes with the service provider well in advance of the class.

Other considerations. The following two considerations can be administratively and legally complex. Conditions might include:

- **Presence of more than one deaf or hard of hearing student in the class.**
  In colleges with large enrollments of deaf and/or hard of hearing students, it is common for two or more of these students to be enrolled in the same class. This does not necessarily mean the same classroom support service(s) are needed by each. This pertains particularly to a situation where one student needs an interpreter and a second student needs real-time speech-to-text services. In this instance, both services should be provided, but presumably the speech-to-text service could

supply notes to both, eliminating the need for a special notetaker.

- **Availability/unavailability of qualified service provider(s)**
  By law, a college cannot conclude that the most appropriate "type" of classroom support service for a given student is unavailable, without clear indication that considerable effort has been made to obtain the services of the needed provider(s). Because of the requisite training factor, one of the CAN systems should be considered among the most available, and a substitute for a steno-based system. The substitution of a transcription system for interpreting depends on several factors mentioned above, including reading proficiency (Brueggemann, 1995).

Automatic Speech Recognition (ASR) in the Classroom

At a national meeting in April 1997 on the topic of "Applications of automatic speech recognition with deaf and hard of hearing people" (Stuckless, 1997), numerous speech scientists spoke enthusiastically about recent developments in the ASR field, with particular reference to the recognition of continuous speech. This coincided with an announcement that Dragon Systems was about to release its first version of NaturallySpeaking, a major product breakthrough (Mandel, 1997). IBM followed later in the same year with ViaVoice.[11]

For many years, scientists have been seeking the model ASR system, one that would have three fundamental properties:[12]

- the capacity to recognize a large vocabulary
- the ability to process natural speech
- the ability to recognize different speakers

Large vocabulary. For more than a decade, systems have been available with vocabularies numbering in the thousands of words. Current products have "active" vocabularies of 30,000 words or more, with the capability of allowing the user to add thousands more, e.g., to add obscure names and technical terms. Vocabulary size *per se* is not a limiting factor for the use of ASR in the college classroom.

Natural speech. Until 1997, commercially available ASR products featured discrete speech recognition, requiring the speaker to pause briefly between each word. While these pauses were tolerable for dictation purposes, speaking in this manner was anything but natural. A secondary effect was that our rate of speech was severely curtailed.

Since 1997, we have been able to choose among a number of products that are capable of recognizing *continuous* speech. By *continuous* we simply mean that no longer must we pause between every word. The provision of continuous speech in ASR certainly enables us to speak more naturally than was possible previously. Also, it enables us to speak at or near our normal speaking rate. A third major advantage is that it tends to lead to greater accuracy, which has been reported as high as 97 percent.

That having been said, we must distinguish between *continuous* and *natural* speech. The two are not synonymous. Continuous speech *per se* does not include the recognition of some of the cues found in natural speech, such as voice inflection and pauses. As a consequence, it does not automatically produce punctuation and other markers, e.g., space between paragraphs, which contribute so much to the readability of text. This is illustrated by the following excerpt from an actual lecture, as transcribed from an audiotape into text, using continuous speech recognition.

> Why do you think we might look at the history of the family history tends to dictate the future okay so there is some connection you're saying what else evolution evolution you're on the right track which changes faster technology or social systems technology

The above excerpt was transcribed with 100 percent verbatim accuracy, using continuous speech recognition. But imagine trying to read lecture text for an hour as it appears above, particularly when it is being displayed at the rate of 150 words per minute. Taken alone, high verbatim accuracy is no guarantee of readability.

As seen next, the same excerpt becomes much more readable when *punctuation* and *speaker identification* are added, using the appropriate voice commands.

---

[11] Both products since have been upgraded and been joined by Lernout and Hauspie's *Voice Xpress* and Philips' *Free Speech*. See Alwang (1998) for a comparative review of these four products.

[12] A recommended clearly-written reference source on ASR is Markowitz, J.A. (1996). *Using speech recognition*. Upper Saddle River, NJ: Prentice Hall.

Instructor: Why do you think we might look at the history of the family?
Student: History tends to dictate the future.
Instructor: Okay. So there is some connection you're saying. What else?
Student: Evolution.
Instructor: Evolution. You're on the right track. Which changes faster, technology or social systems?
Student: Technology.

Recognition of different speakers. A single speaker transcribed the excerpt above because at present, ASR products are incapable of recognizing more than a single speaker (user) at a time, i.e., they lack speaker-independence. To become a user, an individual must sign on and devote half an hour or more to (a) becoming oriented to the system, and (b) orienting the system to his/her distinctive speech characteristics. She/he can then become a user, with her/his own speech files. To use the system, the user identifies her/himself, calling up these speech files.

Without speaker-independent ASR, we cannot pass around a microphone to students in a class with the expectation that their speech will be recognizable. This is one of several reasons why ASR products cannot yet capture conversational speech (Allen, 1997; Woodcock, 1997).

Extending ASR applications into the classroom. Given the present (1999) state of the art, it is not feasible to apply ASR for general real-time classroom use with deaf and hard of hearing students. However, if the application consists of a single user, e.g., a single instructor presenting an uninterrupted lecture, the task becomes less formidable. The following passage was transcribed from an audiotape of another lecture, using ASR.

> Today I'd like to discuss with you a little bit about the history of money my purposes to give you a flavor for the role of money and some of the interesting problems and types of money that existed throughout history to begin with I'd like to raise the question as to where did money come from today how to paper money get here

Note that this monologue is easier to read than the previous unpunctuated passage that involved numerous changes in speakers. Parenthetically, this passage contains two ASR errors (purposes/purpose is; to/did), and a 97% verbatim accuracy rate. Judge its readability for yourself, notwithstanding its absence of punctuation. You may agree that this passage is quite intelligible, in spite of its two ASR transcription errors.

Now let's say the instructor had said *period* or *question mark* as he was speaking to break up his four sentences. These commands not only insert punctuation but also lead automatically to capitalization of the first word in the following sentence, adding to readability. The passage would then have appeared as follows:

> Today I'd like to discuss with you a little bit about the history of money. My purposes to give you a flavor for the role of money and some of the interesting problems and types of money that existed throughout history. To begin with I'd like to raise the question as to where did money come from today. How to paper money get here?

We are not suggesting that the instructor with a class consisting predominantly of hearing students use this strategy, but this sample does suggest how close we have come to making ASR feasible under specific conditions.

One researcher is presently exploring the use of *shadowing* as an interim technique for the use of ASR in the college classroom. This project involves the services of someone with an aptitude for shadowing the speech of the instructor and students together with a few hours of training and practice with ASR.

This person uses a special mask with a built-in microphone connected to a computer containing ASR software and her speech files. Her task is to listen to the instructor, restating what is being spoken as fully as possible, adding sentence-ending punctuation, and identifying each change in speakers, all in real time (Stuckless, in progress).

If recent progress is any indication, there is reason to be optimistic about extending the application of automatic speech recognition into the classroom (Levitt, 1997; Mandel, 1997; Picheny, 1997). Has its time arrived? The answer has to be *no*. However, within a few years, automatic speech recognition is likely to replace other real-time speech-to-text and

notetaking services for many deaf and hard of hearing students in the college classroom. If and when this occurs, it will come about because of its demonstrated value to these students, its relatively low cost, its convenience including availability when needed, and the direct control it will give to the student.

Conclusions

Speech-to-text systems have increased the educators' tools for effectively supporting deaf and hard of hearing students who are educated with hearing classmates. Currently there are many mainstreamed students who cannot hear well enough to follow the classroom discussion, but have intelligible speech and good reading skills. Such students are sometimes given an interpreter, but this service is of limited benefit if the student does not understand signs well.

There are also some situations where the student understands sign communication, but for success in a particular class, it is important after class to be able to review a text that details the class discussion. Speech-to-text services provide a quality option that can effectively address such situations.

The two technologies currently in use to provide speech-to-text services are steno-based systems in which a stenotype machine is linked to a computer, and CAN systems that use standard keyboard laptop computers. Automatic speech recognition systems, in which the conversion to print is done entirely by computer and without an intermediary, will become available in the future and may support communication access even more effectively (Kurzweil, 1999). Other advances in technology are also likely to make these systems more flexible and easier to use.

A serious issue is the fact that none of the speech-to-text technologies discussed in this report adequately address expressive communication by deaf and hard of hearing people.

Individuals with intelligible speech, such as many who are hard of hearing or late deafened, may be able to use their voices to make a comment or ask a question. Others may write or type into a keyboard to produce text or synthetic speech, but in many situations these means may be limited or inadequate. Speech-to-text services are not a panacea for the communication difficulties of deaf and hard of hearing students. In instructional situations such as small group discussions, laboratories, and one-to-one tutoring, these services may be less appropriate than they are in lecture situations (Haydu & Patterson, 1990). Furthermore, many deaf students prefer an interpreter to a speech-to-text system in most class situations (Stinson et al., 1988).

Even with these limitations, speech-to-text services have been used repeatedly to effectively support accessibility to information in the classroom. This experience has clearly demonstrated that these services are a viable option for supporting the communication access of many deaf and hard of hearing students in settings where they are interacting with hearing people. In the future, as the necessary technologies improve, and as we learn more about how these services can effectively support students, speech-to-text services should make even greater contributions to improving the postsecondary education of students who are deaf or hard of hearing.

Postscript pertaining to laws and regulations[13]

With relation to deaf and hard of hearing students, higher education is currently on the horns of a dilemma: given the advent of various speech-to-text systems and advances in voice recognition software, will institutions forego the services of sign language interpreters in reliance on speech-to-text systems, and/or will the shortage of qualified sign language interpreters in certain areas of the country inadvertently push colleges and universities into taking this step?

There are no easy answers. This chapter lays out the pros and cons of various speech-to-text systems and the factors, both student related and instructional, which should enter into a college's determination as to whether speech-to-text is a reasonable accommodation and if so, which type of speech-to-text system would be appropriate in a given circumstance. It also demonstrates that the data suggests that speech-to-text systems can be very effective for a good number of students, but that regardless of future developments, speech-to-text systems will always have the limitations inherent in such a process, most notably, reducing the ability of

deaf and hard of hearing students to fully participate in classes conducted in an interactive manner.

Ultimately, the law requires two things: (a) that communications with students with disabilities, here deaf and hard of hearing students, be "as effective as" that provided to students without disabilities; and (b) that an individualized assessment be made in order to determine what (a) is. This chapter goes a long way toward helping service providers make those assessments. In addition, public colleges and universities must give "primary consideration" to the communication preferences of deaf and hard of hearing students, although as discussed in other commentaries herein, this does not mean the student will always get what s/he wants.

For the most part, if a student prefers sign language and uses interpreters, institutions will opt for providing notes to students via notetaking systems which are effective but less expensive than a speech-to-text system which would arguably provide more complete notes. However, the law does not require that students with disabilities receive the "best" notes, only that they have notes which are "effective." Deaf and hard of hearing students should bear in mind that most hearing students rarely take notes of the quality which would be provided by a speech-to-text system.

At present, speech-to-text systems are roughly as expensive as sign language interpreters. In the future, this may change and lowered costs may become an incentive for institutions to choose speech-to-text over interpreters. Nevertheless, until and unless the law is amended, the legal analysis of which type of auxiliary aid or service should be provided and thus, whether access is achieved, will remain the same.

In addition, if a student's communication preference is speech-to-text and this is not available, the Office for Civil Rights (OCR) has made clear that a good faith effort to locate and implement such a system must be demonstrated before a public institution may provide an alternative system of communication. While private colleges and universities do not have to give "primary consideration" to students' communication preferences, they must nevertheless provide communications which are "as effective as" those provided to students without disabilities. Thus, in order for a private institution to provide an auxiliary aid or service which is arguably less effective than that requested by the student, it should likewise be able to demonstrate that it made a good faith effort to secure the auxiliary aid or service which is "as effective as" that provided nondisabled students, but nevertheless was unable to secure that aid or service.

[1] Contributed by Jo Anne Simon, consultant/attorney specializing in laws and regulations pertaining to students with disabilities.

References

Allen, J. (1997). Applications of automatic speech recognition to natural language and conversational speech. In R. Stuckless (Ed.), *Frank W. Lovejoy symposium on applications of automatic speech recognition with deaf and hard of hearing people*. Rochester, NY: Rochester Institute of Technology.

Alwang, G. (1998). Speech recognition: Finding its voice. *PC Magazine*. October 20 issue.

Brueggemann, B.J. (1995). The coming out of deaf culture and American Sign Language: An exploration into visual rhetoric and literacy. Rhetoric Review: *13*, 409-420.

Cuddihy, A., Fisher, B., Gordon, R. & Schumaker, E. (1994, April).C-Note: A computerized notetaking system for hearing-impaired students in mainstream postsecondary education. *Information and Technology for the Disabled*.1,2.

Eisenberg, S. & Rosen, H. (1996, April). *Real-time in the classroom*. Paper presented at the biennial conference on postsecondary education for persons who are deaf or hard of hearing. Knoxville, Tenn.

Everhart, V.S., Stinson, M.S., McKee, B.G., Henderson, J. & Giles, P. (1996, April). *Evaluation of a speech-to-print transcription system as a resource for mainstreamed deaf students*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Giles, P. (1996). C-Print™ *support service policy*. Unpublished document. National Technical Institute for the Deaf, Rochester, NY.

Hastings, D., Brecklein, K., Cermak, S., Reynolds, R., Rosen, H. & Wilson, J. (1997). *Notetaking for deaf and hard of hearing students*. Report of the National Task Force on Quality of Services in the Postsecondary Education of Deaf and Hard of Hearing Students. Rochester, NY: Northeast Technical Assistance Center, Rochester Institute of Technology.

Haydu, M.L. & Patterson, K. (1990, October). *The captioned classroom: Applications for the hearing-impaired adolescent*. Paper presented at the Fourth National Conference on the Habilitation and Rehabilitation of Hearing Impaired Adolescents. Omaha, NB.

Hobelaid, E. (1988). *The feasibility of using a portable instacap system in a college environment*. ERIC Document 293 614. Canadian Captioning Development Agency, Toronto, Ontario, Canada.

James, V. & Hammersley, M. (1993). Notebook computers as notetakers for handicapped students. *British Journal of Educational Technology. 24*, 63-66.

Kanevsky, D. Nahamoo, D., Walls, T. & Levitt, H. (1992). *Prospects for stenographic and semi-automatic relay service*. Paper presented at meeting of the International Federation of the Hard of Hearing, Tel Aviv, Israel.

Knox-Quinn, C. & Anderson-Inman, L. ((1996, April). *Project CONNECT: An interim report on wirelessly-networked notetaking support for students with disabilities*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Kozma-Spytek, L. & Balcke, M. (1995). Computer-assisted notetaking: Alternative to oral interpreting. *GA-SK Newsletter*, 26, 8.

Kurzweil, R. (1999). *The age of spiritual machines*. New York: Viking Press.

Levitt, H. (1994). Communications technology and assistive hearing technology. In M. Ross ((Ed.) *Communication access for persons with hearing loss*. Baltimore, MD: York Press.

Levitt, H. (1997). Automatic speech recognition: Exploring potential applications for people with hearing loss. In R. Stuckless (Ed.) *Frank W. Lovejoy symposium on applications of automatic speech recognition with deaf and hard of hearing people*. Rochester, NY: Rochester Institute of Technology.

Mandel, M. (1997). Discrete word recognition systems and beyond: Today and five-year projection. In R. Stuckless (Ed.) *Frank W. Lovejoy symposium on applications of automatic speech recognition with deaf and hard of hearing people*. Rochester, NY: Rochester Institute of Technology.

Markowitz, J.A. (1996). *Using speech recognition*. Upper Saddle River: Prentice Hall.

McKee, B.G., Stinson, M.S., Everhart, V.S. & Henderson, J.B. (1995, April). *The* C-Print™ *project: Development and evaluation of a computer-aided speech-to-print transcription system*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

McKee, B.G., Stinson, M., Giles, P., Colwell, J., Hager, A., Nelson-Nasca, M. & MacDonald, A. (1998). C-Print™: *A computerized speech-to-print transcription system.: A guide for implementing* C-Print™. Rochester, NY: National Technical Institute for the Deaf, Rochester Institute of Technology.

Messerly, C. & Youdelman, K. (1994, June). *Computer-assisted notetaking for mainstreamed hearing-impaired high school students*. Paper presented at the International Convention of the Alexander Graham Bell Association for the Deaf, Rochester, NY.

Moore, K., Bolesky, C. & Bervinchak, D. (1994, July). *Assistive technology: Provides equal access in the classroom*. Panel discussion presented at the International Convention of the Alexander Graham Bell Association for the Deaf. Rochester, NY.

Picheny, M. (1997). Continuous speech recognition systems: Today and five-year projection. In R. Stuckless (Ed.) *Frank W. Lovejoy symposium on applications of automatic speech recognition with deaf and hard of hearing people*. Rochester, NY: Rochester Institute of Technology.

Preminger, J. & Levitt, H. (1997). Computer-assisted remote transcription: A tool to aid people who are deaf or hard of hearing in the workplace. *Volta Review*, 99, 219-230.

Sanderson, G., Siple, P. & Lyons, B. (1999). *Interpreting*. Report of the National Task Force on Quality of Services in the Postsecondary Education of Deaf and Hard of Hearing Students. Rochester, NY: Northeast Technical Assistance Center, Rochester Institute of Technology.

Smith, S.B. & Ritterhouse, R.K. (1990). Real-time graphic display: Technology for mainstreaming. *Perspectives in Education and Deafness*. 9(2), 2-5.

Stinson, M.S. & Stuckless, R. (1998). Recent developments in speech-to-print transcription systems for deaf students. In A. Weisel (Ed.). *Issues unresolved: New perspectives on language and deaf education*. Washington, D.C: Gallaudet University Press.

Stinson, M., Stuckless, R., Henderson, J.B. & Miller, L. (1988). Perceptions of hearing-impaired college students toward real-time speech-to-print: RTGD and other support services. *Volta Review*, 90, 336-348.

Stuckless, R. (1983). Real-time transliteration of speech into print for hearing-impaired students in regular classes. *American Annals of the Deaf*. 128, 619-24.

Stuckless, R. (1994). Developments in real-time speech-to-text communication for people with impaired hearing. In M. Ross (1994). *Communication access for people with impaired hearing*. Baltimore, MD: York Press.

Stuckless, R. (Ed.)(1997). *Frank W. Lovejoy symposium on applications of automatic speech recognition with deaf and hard of hearing people*. Rochester, NY: Rochester Institute of Technology.

Task Force on Realtime Reporting in the Classroom (1994). *Realtime in the educational setting: Implementing new technology for access and ADA compliance*. Vienna, VA.:National Court Reporters Foundation.

Virvan, B. (1991). You don't have to hate meetings - Try computer-assisted notetaking. *SHHH Journal*. 12, 25-28.

BEST COPY AVAILABLE

Warick, R. (1994, July). *Campus access for students who are hard of hearing*. Paper presented at the meeting of the Association for Higher Education and Disability, Columbus, OH.

Warick, R., Clark, C., Dancer, J. & Sinclair, S. (1997). *Assistive listening devices*. Report of the National Task Force on Quality of Services in the Postsecondary Education of Deaf and Hard of Hearing Students. Rochester, NY: Northeast Technical Assistance Center, Rochester Institute of Technology.

Woodcock, K. (1997). Ergonomics of applications of automatic speech recognition for deaf and hard of hearing people. In R. Stuckless (Ed.), *Frank W. Lovejoy symposium on applications of automatic speech recognition with deaf and hard of hearing people*. Rochester, NY: Rochester Institute of Technology.

Appendix

Student evaluation of real-time stenographic services[13]

We are attempting to evaluate the effectiveness of our Real-time services to better serve the needs of all students. You can help us assess this program by your honest answers to this questionnaire. Feel free to make any comments you think will help us evaluate this program. All answers will be analyzed by use of a computer to help preserve anonymity and confidentiality. THANK YOU!

Name of Real-time stenotypist _____ Your name (will not be given out) _____ SUBJECT _____

I have had Real-time services at school before (Yes / No)_____ If Yes, where? _____

In class, I read the lecture on the Real-time screen (Yes / No)_____

I get the printed Real-time notes (Yes / No)_____

Circle the responses which best describe your feelings about the stenotypist in this class, **1 being best and 5 being worst**, or NA if not applicable.

My stenotypist:

Displays class information accurately	NA	1	2	3	4	5
Spells most words correctly	NA	1	2	3	4	5
Formulates sentences clearly	NA	1	2	3	4	5
Uses appropriate punctuation	NA	1	2	3	4	5
Identifies different speakers	NA	1	2	3	4	5
Keeps up with the speed of the class	NA	1	2	3	4	5
Seems to concentrate on the task	NA	1	2	3	4	5
Voices for me when requested	NA	1	2	3	4	5
Is comfortable when I ask questions	NA	1	2	3	4	5
Has a friendly attitude	NA	1	2	3	4	5
Stays in stenotypist role	NA	1	2	3	4	5
Is on time to class	NA	1	2	3	4	5
Sets up equipment in a professional manner	NA	1	2	3	4	5
When printed, text is paragraphed and readable	NA	1	2	3	4	5

Comments (use back of sheet if necessary) _____

[13] Adapted from form used at the National Center on Deafness at California State University at Northridge.